

DOCUMENT RESUME

ED 236 612

CS 207 672

AUTHOR Elliott, William R.; And Others
TITLE Measuring the Perceived Reality of Television:
Perceived Plausibility, Perceived Superficiality and
the Degree of Personal Utility.
PUB DATE Aug 83
NOTE 40p.; Paper presented at the Annual Meeting of the
Association for Education in Journalism and Mass
Communication (66th, Corvallis, OR, August 6-9,
1983).
PUB TYPE Reports - Research/Technical (143) --
Speeches/Conference Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS College Students; Content Analysis; Higher Education;
*Measurement Techniques; *Perception; *Programing
(Broadcast); *Research Methodology; Student
Attitudes; *Television; *Television Research; Test
Reliability; Test Validity
IDENTIFIERS *Perceived Reality

ABSTRACT

A tool was devised to measure the perceived reality of television, an important variable in understanding how television influences an individual's perception of social reality. This was accomplished by using a research plan that allowed university subjects to generate the measures used for tapping the "reality" of television content. University students from large introductory telecommunication and speech courses were given a questionnaire that asked them to evaluate the realism of each of 30 television programs, and to explain why they perceived particular programs as realistic and others as unrealistic. In addition they indicated which of five programs they thought most realistic and which least realistic. The data supported three new reality measures: (1) perceived plausibility--the extent to which a program is viewed as providing social and environmental images like those existing in the real world, (2) perceived superficiality--the extent to which a program is viewed as dealing with trivial matters in a repetitious manner, and (3) degree of personal utility--the extent to which the information contained in a program is seen as useful by the viewer. Each of the measures was proven reliable in tests of internal and temporal consistency. Validity was investigated by comparing the correlations between the three scales and an established reality measure, general perceived program reality. This produced support for the conclusion that the tests were concurrently valid. (Extensive tables of data are appended.) (HOD)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED236612

Measuring the Perceived Reality of Television:
Perceived Plausibility, Perceived Superficiality
and the Degree of Personal Utility

by

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as
received from the person or organization
originating it.
☐ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official NIE
position or policy.

William R. Elliott
Department of Humanities and Communications
Drexel University
Philadelphia, PA 19104

Robert Rudd and Leslie Good
Department of Speech
University of Oregon
Eugene, Oregon 97403

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

William R. Elliott

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

Paper Presented to the Theory and Methodology Division
Association for Education in Journalism and Mass Communication
Annual Convention August 6-9, 1983
Corvallis, Oregon

Research on the "cultivation effects" of television has generated one of the liveliest and more public debates about media influence since the Surgeon General's Scientific Advisory Committee's 1972 report on the relationship between television violence and aggressive behavior. Two issues dominate the debate. The first issue, does a cultivation effect even exist, seems the less important, at least for now. The more important issue, and one more consistent with the overall pattern of findings, is what are the conditions under which a cultivation effect might be found?

The fundamental hypothesis of cultivation research is that heavy exposure to the consistent messages presented by television causes audiences to adopt television's characterizations of life.

Cultivation researchers have discovered among many findings that heavy viewers are more likely than light viewers to see the world as mean and violent (Gerbner et al., 1978). Heavy viewers of younger age groups are more likely than light viewers to think of the elderly in

negative terms (Gerbner et al., 1980b).

Recently, the types of analysis used in cultivation studies have come under attack. Hughes (1980) and Hirsch (1980) each re-analyzed data reported by Gerbner et al. (1978) in Violence Profile No. 9. Each, after using multivariate analysis techniques, could find no evidence of a cultivation effect. Others have also failed to locate the cultivation effect (Wober, 1978; Wober, 1982; Skogan & Maxfield, 1980). Some have found the effect limited to certain types of programs (Hawkins & Pingree, 1980) or to specific subsamples of the population (Doob & Macdonald, 1979).

Some of the concerns about cultivation analysis were addressed in Violence Profile No. 11 (Gerbner et al., 1980a). The concept of "mainstreaming" was created to explain the similarity in personal fear of crime indicated by heavy viewers from different demographic groups while light viewers from the same demographic groups differed considerably on the fear of crime variable. "Resonance" was invented to explain what appears to be an amplification effect for cultivation when television images are congruent with the audience's actual life experience. For instance, the difference in fear of crime between heavy and light viewing women is much greater than for men while the difference for heavy and light viewing urban residents is much greater than for suburbanites. Morgan (1983) found that viewers whose TV counterparts are frequently victims show strong relationships between viewing and perceived vulnerability. Bouman (1982) was less successful in demonstrating either a mainstreaming or a resonance effect for Dutch subjects. The problem may be, he

admitted, in the differences in the content of Dutch and American TV. Finally, Wakshlag et al. (1983) used an experimental design to test the relationship between subjects' apprehension about victimization and their reaction to crime programs. Their findings indicated that apprehensive subjects found crime programs more exciting than less apprehensive subjects and that apprehension was positively associated with how intensely subjects evaluated the crime programs. They concluded that the explanation for the relationship between viewing of crime programs and a fear of crime is one based on selective exposure (apprehensive viewers find such programs stimulating) and not on the cultivation hypothesis.

Social reality researchers have also begun to look at variables other than television exposure. Pingree (1978) found that television commercials influenced children's attitudes toward women only when they believed the commercials to be "real." Similarly, Elliott and Slater (1980) found perceived television realism more strongly related to program viewing frequency than to direct experience or gender, suggesting that the perceived realism of television programming rather than mere television exposure might partially explain the relationship between viewing and social reality claimed by cultivation researchers.

Other studies have suggested that a mistake in cultivation analysis may be a reliance on overall television viewing rather than a focus on the specific content viewed. Hawkins and Pingree (1981a) used an Australian sample to test whether differences in the types of television programs produced different cultivation effects. They

found that only specific program types (primarily crime adventure, game shows, and cartoons) were related to perceptions of violence in society. Slater and Elliott (1982) used commonality analysis on data collected from an adolescent sample to investigate the relationships that demographic, experiential, and viewing/reality independent variable sets have on a subject's image of societal safety, understanding of law enforcement, and acceptance of the way television portrays police behavior. Of the three independent variable sets, the greatest proportion of variation in each of the three dependent variables was explained by the viewing/reality set. The three variables making up this set--total television viewing, law enforcement program viewing, perceived program realism--were used in a second commonality analysis to determine which of them explained the greatest proportion of variation in the dependent variables. Perceived program realism was the most important independent variable (making the greatest unique contribution to variation explained) while law enforcement program viewing was second most important. The independent contribution of total television viewing to variation explained was minimal for any of the dependent variables.

After an extensive review of the research on television and social reality, Hawkins and Pingree (1981b) made several important suggestions. They noted that researchers bring their own vision of what television is to the research setting. To the degree that the researchers' visions differ from most audience members, any media influence on reality perception may be obscured. They also pointed out that most of the measurement used in social reality research is

crude and poorly conceptualized. In future work they suggested that greater concern with measurement regarding the ways individuals use television to form reality perceptions be practiced.

This paper is concerned with defining and measuring the perceived reality of television programming. Much of the work done thus far has viewed the perceived reality of television as a univariate concept. Elliott and Slater (1980) asked respondents to evaluate television programs along a five point scale ranging from extremely realistic to extremely unrealistic. Later they summed across six such measures to form a "perceived law enforcement program realism" scale (Slater & Elliott, 1982). Pingree (1978) used the factor score from a set of six general questions dealing with the realism of television programming and Greenberg (1974) reported using a three item scale to measure the "perceived reality of television" although the three items he used were not made explicit in the article.

Other efforts have stressed the multidimensional nature of the perceived reality of television programming. Hawkins (1977) used factor analysis to identify two major dimensions of children's perceptions of television reality: the magic window, "the degree to which television is seen as portraying real life instead of fiction" (p. 311) and the social expectations dimension, "the degree to which television's characters and events are similar to children's expectations about the real world" (p. 311). Ostman and Jeffers (1981) were able to successfully apply these child derived dimensions to a study comparing schizophrenic-labeled adults with a "normal"

adult sample. Reeves (1978) also used factor analysis to investigate children's perceptions of television reality. He found three of his items clustering along a prosocial reality factor and two items forming an antisocial factor.

These studies attempting to measure social reality have several problems. First, they have tended to rely on conceptualizations of the perceived reality of television largely formulated by the researchers and not by the subjects. Second, in some cases the number of items forming the scales used has been extremely small (for example: one item in the case of Elliott & Slater, 1980; two items for Reeves' antisocial factor and three items for his prosocial factor, 1978; three items in the case of Greenberg, 1974). Third, and perhaps most important, the authors of these studies have generally failed to provide any information on scale reliability and validity. Such an omission makes adoption of any such measuring instruments chancy. Because perceived television reality is an important variable in understanding how television influences an individual's perception of reality (Slater & Elliott, 1982; Lagerspetz & Viemerö, 1982; Comstock et al., 1978), we decided to develop a measurement tool for the perceived reality of television.

To accomplish this we adopted a research plan that allowed our subjects to generate the measures used for tapping the "reality" of television content. In addition, we used several waves of data collection to provide replication, allow reliability estimates, and to gather information on concurrent and predictive validity. Finally, our sample was composed of young adults and not children

giving our findings, perhaps, greater application for use among adolescent or adult samples.

METHOD

Item Generation

All of the subjects used during the study were university students recruited from large introductory lecture classes in telecommunication and from introductory speech courses.

In early February of 1982 we selected a nonrandom sample of ten students to perform a Q-sort on 121 television programs that had appeared from the 1974-1975 season until the time of the questionnaire.¹ The subjects sorted cards bearing the names of the 121 programs into nine piles ranging from the most realistic programs (given the value of 9) to least realistic (with the value of 1). The students were instructed to place the unfamiliar programs in the neutral (5) category. After the Q-sort, we asked the students why they had rated specific programs as realistic or unrealistic. Responses were tallied, collated, and listed as potential questionnaire items.

We included in a questionnaire thirty programs whose Q-sort scores ranged from realistic to unrealistic. The questionnaire asked subjects to evaluate each program's realism and to explain why they perceived particular programs as realistic and others as

unrealistic. One hundred and two subjects provided us with a rough measure of perceived program realism for each of the 30 shows. Subjects in this part of the study checked whether they saw the program as extremely realistic, slightly realistic, neutral, slightly unrealistic, or extremely unrealistic. If they were completely unfamiliar with a program they left that response blank.

In addition, the subjects indicated which of five programs (Quincy, Barney Miller, Taxi, Hill Street Blues, M*A*S*H) they thought most realistic and which of five programs (The Love Boat, The Dukes of Hazzard, Mork & Mindy, The Greatest American Hero, Three's Company) they thought least realistic. ² Hill Street Blues was seen as most realistic program by 45.1% of the sample. M*A*S*H ranked second with 27.5%. The least realistic program was Mork & Mindy, at 45.1% of the sample, followed, distantly, by The Dukes of Hazzard at 16.7%. After selecting the most realistic and unrealistic programs, the subjects were asked to list up to three reasons why they chose those programs.

The reasons were tabulated, collated, and added to our pool of possible items and topics for inclusion in the reality scale. The new items, those collected during the Q-sort interviews, several selected from the works of others, and our own ideas were combined into a single source pool of possible questions. Forty statements regarding the reality of television programs were thus developed.

Scale Development

Time One. A questionnaire designed to determine scale inclusion or exclusion was administered in March of 1982 to a sample of 246 college students. The questionnaire included an overall measure of perceived realism for 20 programs³ (the same measure used in the previous questionnaire) and the 40 existing items in the source pool of possible questions. Students were asked to apply the 40 source pool questions either to M*A*S*H (N=126) or to Happy Days (N=120). Questionnaires directed at each program were alternated as they were passed out to the students in class. M*A*S*H and Happy Days were selected as stimuli for this portion of the study because the students had previously said that they were familiar with both programs (101 of 102 subjects in the previous survey indicated familiarity with M*A*S*H while 97 of 102 subjects were familiar with Happy Days). Also, 63% of the earlier sample indicated that they thought M*A*S*H slightly or extremely realistic while 57.8% of the sample indicated that Happy Days was either slightly or extremely unrealistic. We wanted to ensure the programs we had picked as stimuli were both familiar to the subjects and different in terms of some general measure of program realism. Hopefully, this would guarantee variability in terms of the responses to the items we had developed.

After performing factor and scale analyses on the time one results of the scale development program, we were left with a pool of 20 usable items. We revised several items in the original pool and

added a number of new items designed to match the dimensions suggested by the factor and scale analyses. A total of 35 items, 15 new plus the 20 from time one, were available for administration to a new set of subjects.

Time Two. A total of 264 students participated in this part of the study. Approximately one-half of the subjects (N=134) received questionnaires with items concerning Dallas while the other half (N=130) responded to questions about Quincy. We selected Dallas and Quincy for the same reasons we used M*A*S*H and Happy Days. Dallas was familiar to 84.2% of the time one sample while Quincy had a familiarity percentage of 90.2. Only 32.3% of the time one sample had indicated that Dallas was extremely or slightly realistic while the comparable figure for Quincy was 54.9%. These two programs were also selected because they were one hour dramatic programs while both M*A*S*H and Happy Days were one-half hour situation comedies.

The results of the time two questionnaires were subjected to a scale analysis and a total of 31 items remained for reliability analysis and for use in the final phase of the study.

Experiment

In November of 1982, a total of 51 volunteer students from introductory speech courses agreed to view a videotape of one of four new television programs. New television programs were selected to improve the likelihood that students would respond to the programs directly rather than to preconceived notions about the programs based on earlier viewing. The programs included one episode of PBS's

offering of Oppenheimer and three NBC offerings: Cheers, Knight Rider and Silver Spoons. Each student participating in the experiment was randomly assigned to one of the programs. After viewing the program, the students completed questionnaires containing the 31 scale items and one general reality question about the program asking each student to evaluate the program they had watched as extremely realistic, slightly realistic, neutral, slightly unrealistic, or extremely unrealistic.

RESULTS

Scale Development and Reliability

Table 1 presents the factor pattern matrix for a principal factor analysis using oblique rotation⁴ after eliminating those items with low factor loadings (approximately .40), loadings of .40 or higher on two or more factors, two or fewer measures on a single factor (Wimmer & Dominick, 1983), or uninterpretable factors. After eliminating these items and running the factor analysis the three resulting identifiable factors were labeled perceived plausibility, perceived superficiality and the degree of personal utility. The actual items used in the development of each factor can be seen in Tables 2, 3 and 4. We used oblique rotation assuming that the extracted factors would very likely be correlated, given the phenomena under investigation.

Table 1 About Here

The time one measures P1 to P10 (specific questions used for these measures can be seen in Table 2) show that the students used for this analysis addressed the question of how closely the program content matched their own life experience. The comparison the students seemed to be making was how closely did the events, characters, and settings used in television programs match the events, people and settings in the "real world." We labeled this dimension as perceived plausibility. Measures S1 to S7 (with questions listed in Table 3) make up the second factor. Here the responses seemed to indicate a concern with trivial television content and with a shallow and repetitious treatment of human problems. We labeled this dimension as perceived superficiality. The third and final factor, made up of measures U1 to U3 (questions listed in Table 4) seemed to deal with the way the students applied the messages from the program to their personal lives. We labeled this dimension as personal utility. The definitions for each of the three factors are listed below:

*Perceived Plausibility = the extent to which a program is viewed as providing social and environmental images like those existing

in the real world;

Perceived Superficiality = the extent to which a program is viewed as dealing with trivial matters in a repetitious manner;

Degree of Personal Utility = the extent to which the information contained in a program is seen as useful by the viewer.

It is interesting to note that two of these factors resemble those discovered by Hawkins (1977). His magic window is very similar to our perceived plausibility while his social expectations parallels our degree of personal utility. Only our perceived superficiality has no counterpart in his research. This may be because we stressed to our subjects that they were to consider what makes a program unrealistic as well as what makes it realistic.

Table 2 About Here

The results of the scale analysis for the perceived plausibility dimension are presented in Table 2. The internal consistency of the items has been evaluated using Cronbach's alpha (Carmines & Zeller, 1979). Only items with item-total correlations of .40 or better were included in the table for analysis. The alphas of .87 associated with both time one and time two indicate that our measures are

internally consistent for perceived plausibility.

Table 3 provides the same basic analysis for the perceived superficiality scale. One of the items that seemed to scale appropriately during time one failed to make the established criteria for inclusion of .40 during time two (S7). As with the alphas of the perceived plausibility scale, the alphas here are quite high, .85 for time one and .83 for time two, indicating, once again, high levels of internal consistency within the measure.

Table 3 About Here

Table 4 shows the internal consistency figures for the degree of personal utility measure. The addition of six new items for the scale increased the alpha from .80 during time one to .87 at time two. The three variables used in the time one analysis were each strongly related to the scale development during time two.

Table 4 About Here

Test-retest correlation figures as well as the intercorrelations

between the three scales are presented in Table 5. All correlations above the diagonal in the table were collected from the subjects who used Dallas as their program while those below the diagonal come from the subjects who used Quincy. The correlations along the diagonal are the test-retest correlations, ranging from a low of .78 for perceived superficiality to .86 for perceived plausibility and .88 for the degree of personal utility. The administration of the test-retest correlation questionnaires occurred after a two week interval. Thirty students provided the data for this part of the analysis. The results of the test-retest analysis agree quite closely with the internal consistency analysis indicated by Cronbach's alpha.

Table 5 About Here

The other correlations in the table show how the three scales are, in fact, intercorrelated for each of the programs used as stimuli in the time two administration. For both programs perceived plausibility shows a strong positive relationship to the degree of personal utility. For Dallas the correlation is .63 and for Quincy the correlation is .54. Perceived plausibility is negatively related to perceived superficiality for both programs (Dallas = $-.69$; Quincy = $-.62$) while perceived superficiality is also negatively related to the degree of personal utility (Dallas = $-.52$; Quincy = $-.29$).

Overall, the pattern of strong reliability estimates provides some certainty that these three scales are consistent both internally and over time.

Validity Measures

The data collected in this study were used to test two types of validity--concurrent and predictive. To test concurrent validity it is necessary to have a related measure which is collected at the same time. If the newly developed measure is valid, it will correlate with the related measure in the predicted direction. Such a related measure was available for this study by using the general perceived program realism measure--the five point response evaluating each listed program along a continuum from extremely realistic (5) to extremely unrealistic (1). Table 6 displays these correlations for the time two subjects and for the experimental subjects.

Table 6 About Here

The correlations in Table 6 are substantial and in the appropriate direction for both sets of subjects. Perceived plausibility is positively related to general perceived program realism, as one would predict ($r = .66$ for time two and $.72$ for the experimentals) and the degree of personal utility is also positively

related to this measure ($r = .51$ for time two and $.66$ for experimentals). Appropriately, the correlations between perceived superficiality and general perceived program realism are negative for both samples, $-.51$ for time two and $-.70$ for the experimentals.

The evaluation of predictive validity depends on the development of measures in one study which are used to forecast the results of a later study. Our procedure for looking at predictive validity was to develop an experimental condition where each of the three scales could be evaluated using a new and unfamiliar (to the subjects) set of stimuli. The appropriate statistic for this type of analysis, resulting from a multivariate analysis of variance, is Wilks' lambda (Kerlinger & Pedhazur, 1973). To test the null hypothesis (that the mean vectors for each of the three dependent variables were equal) we had our experimental subjects watch one of four new programs on television--Oppenheimer, Cheers, Knight Rider and Silver Spoons.

Table 7 About Here

Results of the multivariate analysis of variance on the three dependent variables, perceived plausibility, perceived superficiality, and the degree of personal utility, are displayed in Table 7. The Wilks' lambda ($.329$) for this analysis is significant at the $p < .001$ level (F-ratio approximation = 5.49 with 9 and 85.33).

degrees of freedom). Consequently it is reasonable to conclude that the hypothesis of equal population mean vectors should be rejected (Huck et al., 1974). In comparing entries for each of the dependent variables in Table 7 (each of the new scales), one can see where many of the differences occur.

Somewhat suprisingly, perhaps, is the observation that the highest mean score for perceived palusibility is associated with Cheers, a situation comedy set in a Boston bar, and not with Oppenheimer, a nonfiction dramatic program. An investigation of the 95% confidence intervals associated with each program on perceived plausibility shows that the confidence interval for Knight Rider falls outside of the intervals for both Cheers and Oppenheimer while Cheers also falls outside of the interval for Silver Spoons. Both Cheers and Oppenheimer seemed to contain information people perceive as true-to-life. The same cannot be said for Silver Spoons, and especially not for Knight Rider.

The results for perceived superficiality are nearly opposite those for perceived plausibility. Knight Rider is perceived as the most superficial of the four programs followed by Cheers (highest on perceived plausibility), Silver Spoons, and Oppenheimer. The 95% confidence intervals show that Knight Rider falls outside of the intervals for Oppenheimer and Silver Spoons, but that there is a slight overlap between Knight Rider and Cheers. No large differences can be detected between the other three shows. Clearly, Knight Rider represents a deviant program, when compared to the other three, concerning the perceived superficiality variable.

The last variable presented in this study is the degree of personal utility. Again, as with perceived superficiality, the unusual program is Knight Rider. The mean scores indicate the highest degree of personal utility for Oppenheimer followed by Silver Spoons and Cheers. The difference between these three programs is minimal, however, with a 23.8 mean score for Oppenheimer, a 22.5 mean score for Silver Spoons, and a 22.2 mean score for Cheers. Compare all of these to Knight Rider's mean of 15.0. Knight Rider's confidence interval falls outside of the intervals for both Oppenheimer and Cheers and nearly so for Silver Spoons.

The differences in the way the mean scores of the programs rank on the three scales indicate each scale is measuring something different. Cheers is ranked first on perceived plausibility, second on perceived superficiality, and third on the degree of personal utility. Oppenheimer has the second highest mean score on perceived plausibility, ranks fourth on perceived superficiality, and first on the degree of personal utility. Silver Spoons has the third highest perceived plausibility score, the third highest mean for perceived superficiality, and the second highest mean score for the degree of personal utility. Only Knight Rider showed consistency across the three scales, ranking fourth on perceived plausibility, first on perceived superficiality, and fourth on the degree of personal utility. These differences indicate that one dimension may emerge as dominant in terms of evaluating a program as realistic or unrealistic.

All of the validity analyses have indicated that the three

scales appear to have concurrent and predictive validity. All correlations in the concurrent validity analysis were statistically significant in the proper direction. In terms of predictive validity, the overall measure of difference in mean vectors for the three scales indicated by Wilks' lambda was significant. An analysis of the means and confidence intervals clearly pointed to the ability of the three measures to differentiate among the four programs used in the experiment. Unfortunately, no direct test for construct validity was available. In the future such a test might be developed to tie perceived plausibility, perceived superficiality, and the degree of personal utility to cultivation analysis-like variables. A very useful test could be made linking these new measures to well documented measures of conflict resolution such as those developed by Roloff and Greenberg (1979).

CONCLUSION

This study began with a discussion of cultivation analysis, moved to questions dealing with the perceived reality of television programming in the shaping of viewer reality perceptions, and ended with the presentation of data supporting three new reality measures labeled: perceived plausibility, perceived superficiality and the degree of personal utility. The scales seem to measure both reality and unreality. This new emphasis on programming as unrealistic was particularly evident in the analysis of the new television program

Knight Rider.

Each of the measures survived tests of their reliability through tests of internal consistency (Cronbach's alpha) and temporal consistency (test-retest correlations). Validity was investigated by comparing the correlations between the three scales and an established reality measure, general perceived program reality. This test of concurrent validity produced support for the conclusion that the tests were concurrently valid.

An experiment was set up to test the predictive validity of the measures. Four new television programs were shown to randomly selected audiences. In each case the three scales performed well in differentiating between the four stimulus programs. As with concurrent validity, the test supported an interpretation of high predictive validity. No test of construct validity was made.

However, in spite of the fact that the scale development proceeded smoothly, problems remain. A total of 31 different items make up the three scales. If a researcher was interested in investigating the effects of each of the scales applied to three television shows (a quite likely circumstance if the study were of hospital shows) 93 items would be required to estimate the realism for the three programs. It might be possible to avoid some of the problems of this type by applying the scales directly to different media (a perceived plausibility for television programs or films) or to specific classes of programs (a perceived superficiality for television family programs) but such extensions have not been evaluated and they, too, would need careful, and time consuming,

development. If there is an interest in reliable and valid measures, then it may be necessary to reduce the scope of what we typically try to accomplish in so many communication studies.

Measurement and conceptualization remain central issues in the social sciences (Blalock, 1982) but the concern does not seem to have fully penetrated our discipline. This paper represents one effort to begin to redress the imbalance favoring data collection and analysis over measurement. Whether or not the social sciences can ever approximate other sciences, such as physics, is certainly an arguable point. Even whether they should aspire to the same level of measurement precision is debatable. Regardless of whether communication studies evolve toward physics or chemistry, it does seem certain that unless measurement is taken more seriously, we will be unlikely to learn much about any of the communication processes, including how media influence reality perceptions. Until we begin to use similar measures for similar concepts we will confront a number of little studies, each offering a tantalizing glimpse of what might, or might not, be.

NOTES

1. Programs were selected for inclusion if they had been in the 20 top rated shows for a season or if they lasted on network television for more than one season. Most of the programs came from Total Television: A Comprehensive Guide to Programming from 1948 to 1980 (McNeil, 1980).
2. The five realistic programs were selected on the basis of the results from the Q-sort. All programs so selected had mean scores above the neutral position of 5.0. Unrealistic programs were similarly selected. These programs all had mean scores below the 5.0 neutral position.
3. The programs included Mork & Mindy, Quincy, Three's Company, The Love Boat, Mary Tyler Moore, Barney Miller, CHIPS, Emergency!, Happy Days, Lou Grant, LaVerne and Shirley, M*A*S*H, Taxi, The Jeffersons, One Day at a Time, Dallas, Starsky and Hutch, The Dukes of Hazzard, Hill Street Blues, and WKRP in Cincinnati.
4. All statistical analyses were run using the SPSS: Statistical Package for the Social Sciences (Nie et al., 1975).
5. Time one subjects were not included since they received only a subset of the items making up the final scale. The scales administered to the Time two and experimental subjects were

exactly the same.

6. The decimal degrees of freedom in multivariate analysis of variance can occur when three or more dependent variables are involved and when there are four or more levels for the main effect (Huck et al., 1974).

BIBLIOGRAPHY

- Blalock, H. M. Jr. (1982). Conceptualization and Measurement in the Social Sciences. Beverly Hills, Calif.: Sage.
- Bouwman, H. (1982). "A replication of Gerbner's research on cultural indicators: the Dutch situation." Paper presented at the Symposium on Cultural Indicators for the Comparative Study of Culture. Vienna, Austria.
- Carmines, E. G. and R. A. Zeller (1979). Reliability and Validity Assessment. Beverly Hills, Calif.: Sage.
- Comstock, G., S. Chaffee, N. Katzman, M. McCombs, and D. Roberts (1978). Television and Human Behavior. New York: Columbia University Press.
- Doob, A. N. and G. E. Macdonald (1979). "Television viewing and fear of victimization: is the relationship causal?" Journal of Personality and Social Psychology 37: 170-179.
- Elliott, W. R. and D. Slater (1980). "Exposure, experience and perceived TV reality for adolescents." Journalism Quarterly 57: 409-414, 431.
- Gerbner, G., L. Gross, M. Jackson-Beeck, S. Jeffries-Fox, and N.

- Signorielli (1978). "Cultural Indicators: violence profile no. 9." *Journal of Communication* 28 (3): 176-207.
- Gerbner, G., L. Gross, M. Morgan, and N. Signorielli (1980a). "The 'mainstreaming' of America: violence profile no. 11." *Journal of Communication* 30 (3): 10-29.
- Gerbner, G., L. Gross, N. Signorielli, and M. Morgan (1980b). "Aging with television: images on television drama and conceptions of social reality." *Journal of Communication* 30 (1): 37-47.
- Greenberg, B. S. (1974). "Gratifications of television viewing and their correlates for British children," in J. G. Blumler and E. Katz (eds.) *The Uses of Mass Communication*. Beverly Hills, Calif.: Sage.
- Hawkins, R. P. (1977). "The dimensional structure of children's perceptions of television reality." *Communication Research* 4: 299-320.
- Hawkins, R. P. and S. Pingree (1980). "Some processes in the cultivation effect." *Communication Research* 7: 193-226.
- Hawkins, R. P. and S. Pingree (1981a). "Uniform messages and habitual viewing: unnecessary assumptions in social reality effects." *Human Communication Research* 7 (4): 291-301.
- Hawkins, R. P. and S. Pingree (1981b). "Using television to construct social reality." *Journal of Broadcasting* 25: 347-364.

Hirsch, P. M. (1980). "The 'scary world' of the nonviewer and other anomalies." *Communication Research* 7: 403-456.

Huck, S. W., U. H. Cormier, and W. G. Bounds, Jr. (1974). *Reading Statistics and Research*. New York: Harper and Row.

Hughes, M. (1980). "The fruits of cultivation analysis: a reexamination of some effects of watching television." *Public Opinion Quarterly* 44: 281-302.

Kerlinger, F. N. and E. J. Pedhazur (1973). *Multiple Regression in Behavioral Research*. New York: Holt, Rinehart and Winston.

Lagerspetz, K. and V. Viemero (1982). "An international investigation of televised violence and aggression in children." *Nordicom Review of Nordic Mass Communication Research* 2.

McNeil, A. (1980). *Total Television: A Comprehensive Guide to Programming from 1948 to 1980*.

Morgan, M. (1983). "Symbolic victimization and real world fear." *Human Communication Research* 9 (2): 146-157.

Nie, N. H., C. H. Hull, J. G. Jenkins, K. Steinbrenner, and D. H. Bent (1975). *SPSS: Statistical Package for the Social Sciences*. Second Edition. New York: McGraw-Hill.

Ostman, R. E. and D. W. Jeffers (1981). "TV police and family programs as reality to schizophrenic labeled persons." *Journalism Quarterly* 58: 65-68, 88.

- Pingree, S. (1978). "The effects of nonsexist television commercials and perceptions of reality on children's attitudes about women." *Psychology of Women Quarterly* 2: 262-277.
- Reeves, B. (1978). "Perceived TV reality as a predictor of children's social behavior." *Journalism Quarterly* 55: 682-689, 695.
- Roloff, M. E. and B. S. Greenberg (1979). "Sex differences in choice of modes of conflict resolution in real-life and television." *Communication Quarterly* 27 (3): 3-12.
- Skogan, W. G. and M. G. Maxfield (1980). *Coping with Crime: Victimization, Fear and Reactions to Crime in Three American Cities*. Evanston, Ill.: Northwestern University Center for Urban Affairs.
- Slater, D. and W. R. Elliott (1982). "Television's influence on social reality." *Quarterly Journal of Speech* 68: 69-79.
- Wakshlag, J. J., L. Bart, J. Dudley, G. Groth, J. McCutcheon, C. Rolla (1983). "Viewer apprehension about victimization and crime drama programs." *Communication Research* 10: 195-217.
- Wimmer, R. D. and J. R. Dominick (1983). *Mass Media Research: An Introduction*. Belmont, Calif.: Wadsworth.
- Wober, J. M. (1978). "Television violence and paranoid perception: the view from Great Britain." *Public Opinion Quarterly* 42: 315-321.

Wober, J. M. (1982). "Where failure is success: the shadow of Gerbner's prophecy does not fall upon Britain." Paper presented at the Symposium on Cultural Indicators for the Comparative Study of Culture. Vienna, Austria.

Table 1

Factor Pattern Loadings After Oblique Rotation
for Measures of the Dimensions of the Perceived Reality
of M*A*S*H (N=126) and Happy Days (N=120)

Factor Pattern Loadings			
Measures	Perceived Plausibility	Perceived Superficiality	Personal Utility
P1	.64	.08	.13
P2	.64	-.10	.00
P3	-.61	.14	-.05
P4	.61	.16	.14
P5	-.61	-.04	-.08
P6	.50	-.22	.13
P7	.45	-.04	-.10
P8	-.43	.11	-.30
P9	.42	-.23	.07
P10	.38	-.23	.03
S1	-.09	.67	-.04
S2	-.19	.67	.01
S3	.11	.60	.00
S4	.12	-.59	.17
S5	.07	.55	-.14
S6	-.22	.51	-.04
S7	-.22	.40	-.06
U1	-.03	-.03	.69
U2	.06	-.14	.68
U3	.16	-.08	.63
Eigenvalue	7.64	1.03	0.57
% of Variation	82.8%	11.1%	6.1%

Table 2

Scale Analysis of Perceived Plausibility
for Time One and Time Two

Item	Item-Total Correlations	
	Time One	Time Two
P1. The places where people in this program live and work are very true-to-life.	.65	.65
P2. The dialogue on (Program) is realistic and believable.	.63	.61
P3. The setting for (Program) just doesn't seem real.*	.66	.60
P4. The surroundings used for this program--buildings, scenery, offices, houses--are really very similar to what I think it would actually be like if I were there in person.	.55	.42
P5. I don't think anyone really lives or works in places like the ones shown on (Program).*	.58	.43
P6. People in this show are like people you or I might actually know.	.69	.65
P8. The way people really live their everyday lives is not portrayed very accurately on this program.*	.63	.65
P9. Events that actually have happened or could happen are shown on this program.	.56	.62
P10. The characters in (Program) look like people you might see every day.	.52	.61

Table 2 (continued)

<u>Item</u>	<u>Item-Total</u> <u>Correlations</u>	
	<u>Time</u> <u>One</u>	<u>Time</u> <u>Two</u>
P11. This program shows that people have both good and bad sides.		.42
P12. I have a hard time believing the characters in (Program) are real because the basic situation is so far-fetched.*		.58
P13. This program deals with the kind of very difficult choices people in real life have to make.		.58
<u>Standardized Alpha Coefficient</u>	.87	.87
(N)	(209)	(160)

Items with an asterisk ("*") are reverse coded. Item P7 was deleted from the analysis because its item-total correlation fell below .40. Items P11, P12, and P13 were not used at Time One but were added for Time Two.

Table 3
Scale Analysis for Perceived Superficiality
for Time One and Time Two

Item	Item-Total Correlations	
	Time One	Time Two
S1. (Program) doesn't present issues seriously.	.67	.58
S2. The feelings portrayed by the actors in this program are too shallow and one-dimensional for the characters to seem real.	.71	.59
S3. Sometimes there are so many things happening and so many different stories going on in this show that it doesn't make any sense.	.48	.42
S4. This show deals with important personal and social ideas.*	.69	.54
S5. Frequently, this particular program treats unimportant matters much too seriously.	.54	.54
S6. This show never changes from week to week.	.62	.47
S7. People in this program stay the same from week to week; they never seem to grow or change.	.53	
S8. There are many times when (Program) deals with the same problems you read about in the newspaper or see on television news.*		.46
S9. Human feelings are ignored on this show so time can be spent on action and excitement.		.51
S10. Real people simply aren't as stupid as the characters on this program.		.54

TABLE 3 (continued)

Item	Item-Total Correlations	
	Time One	Time Two
S11. (Program) doesn't have any plot or real meaning.		.54
<u>Standardized Alpha Coefficient</u>	<u>.85</u>	<u>.83</u>
(N)	(209)	(160)

Items with an asterisk ("*") are reverse coded. Item S7 was deleted from the Time Two analysis because its item-total correlation fell below .40. Items S8, S9, S10, and S11 were not used during Time One but were added for Time Two.

Table 4

Scale Analysis of the Degree of Personal Utility
for Time One and Time Two

Item	Item-Total Correlations	
	Time One	Time Two
U1. I use the ideas from this show when I find myself in similar real-life situations.	.60	.71
U2. The characters in (Program) give me a lot of insight into myself.	.69	.64
U3. I've learned a lot about the kinds of work people do in the jobs they have on this program.	.68	.57
U4. This program helps me to understand what life would be like if I lived someplace different than where I live now.		.64
U5. I think this program has helped me to understand the type of person I would like to be.		.55
U6. The ideas on this show about how to deal with people are helpful to me when I have to work with people in my own life.		.71
U7. I think the values they stress in (Program) are the type that would be good for most people to follow.		.51
U8. One of the things I like about this show is that I get to see how people from other parts of the country live.		.45
U9. This program is very good for teaching me about how people in other parts of the country and world live.		.67
Standardized Alpha Coefficient	.80	.87
(N)	(209)	(160)

Items U4, U5, U6, U7, U8, and U9 were not used for Time One but were added for Time Two.

Table 5
Time Two Intercorrelations Between Reality Measures
and Test-Retest Correlations (Along Diagonal)

	Perceived Plausibility	Perceived Superficiality	Degree of Personal Utility
Perceived Plausibility	.86	-.69	.63
Perceived Superficiality	-.62	.79	-.52
Degree of Personal Utility	.54	-.29	.88

The intercorrelations above the diagonal are for the measures collected for Dallas (N=80) while those below the diagonal are for Quincy (N=82). The figures along the diagonal are the test-retest correlations for each measure (N=30).

Table 6

Intercorrelations Between General Perceived Program Realism,
Perceived Plausibility, Perceived Superficiality, and the
Degree of Personal Utility for Time Two and Experimental Subjects

Correlations Between General Perceived Program Realism and:	Time Two Subjects	Experimental Subjects
Perceived Plausibility	.66	.72
Perceived Superficiality	-.51	-.70
Degree of Personal Utility	.51	.66
(N)	(159)	(40)

Table 7

Means, Standard Deviations and 95% Confidence Intervals
for Perceived Reality Measures by Experimental Treatment

Perceived Reality Measure by Treatment	Mean	Standard Deviation	N	95% Confidence Interval
Perceived Plausibility				
<u>Cheers</u>	45.8	6.9	11	41.2 to 50.5
<u>Oppenheimer</u>	42.3	9.6	12	36.2 to 48.5
<u>Silver Spoons</u>	35.9	5.7	8	31.1 to 40.5
<u>Knight Rider</u>	28.9	8.2	10	23.0 to 34.8
Total Sample	38.7	10.1	41	35.5 to 41.9
Perceived Superficiality				
<u>Knight Rider</u>	32.6	4.1	10	29.7 to 35.5
<u>Cheers</u>	26.9	4.8	11	23.7 to 30.1
<u>Silver Spoons</u>	25.1	3.7	8	22.0 to 28.2
<u>Oppenheimer</u>	23.8	4.5	12	21.0 to 26.7
Total Sample	27.0	5.4	41	25.4 to 28.7
Degree of Personal Utility				
<u>Oppenheimer</u>	23.8	3.9	12	21.3 to 26.3
<u>Silver Spoons</u>	22.5	5.2	8	18.1 to 26.9
<u>Cheers</u>	22.2	5.7	11	18.3 to 26.0
<u>Knight Rider</u>	15.0	4.5	10	11.8 to 18.2
Total Sample	21.0	5.8	41	19.1 to 22.8

The F approximation for Wilks' Lambda ($\Lambda=.329$) is 5.49 with 9 and 85.33 degrees of freedom which is significant at the $p<.001$ level.